Robot Learning Beyond Imitation: Opportunities and Validation Challenges

How is robot learning moving beyond imitation learning, and what validation challenges must be solved before it can be trusted in the real world?

1st Alireza Safdari Khosroshahi
TU Berlin
Berlin, Germany
alireza.safdarikhosroshahi@campus.tu-berlin.de

***Abstract*—Robot learning has become a key paradigm for enabling robots to operate in complex, unstructured, and contact-rich environments where manual programming is insufficient. While imitation learning has long provided an efficient way to transfer human expertise to robotic systems, its dependence on demonstration quality and limited robustness under distribution shift restrict its scalability in open-world settings. Recent advances therefore move beyond classical imitation toward reinforcement learning, diffusion-based policy learning, and foundation-model-based robotics, aiming to improve autonomy, generalization, and task diversity. These approaches have significantly expanded the capabilities of robotic systems, particularly in manipulation and embodied decision-making, but they also introduce new challenges related to safety, interpretability, embodiment mismatch, and reliable real-world deployment [1]–[4]. This report surveys the progression from imitation-based methods to more general robot learning paradigms and argues that validation has emerged as a central bottleneck for future progress. In particular, the growing complexity and data dependence of learned robotic policies demand more rigorous approaches to testing, scenario-based evaluation, and trustworthiness assurance. Overall, the report highlights that advances in robot learning must be accompanied by equally strong advances in validation methodology if such systems are to achieve robust and trustworthy deployment beyond laboratory environments.**

***Index Terms*—foundation models, imitation learning, reinforcement learning, robot learning, robot manipulation, safety assurance, trustworthiness, validation**

I. Introduction

Robot learning has become a central research direction in modern robotics because many real-world tasks are too diverse, unstructured, and contact-rich to be solved reliably through manual programming alone. In manipulation and service robotics, a robot must cope with changing object configurations, uncertainty, and partially specified goals, making purely hand-engineered pipelines difficult to scale [5]. For this reason, learning-based methods have increasingly replaced rigid task programming, with imitation learning emerging as one of the earliest and most practical approaches for transferring expert skills to robotic systems. More recently, the field has expanded beyond imitation toward reinforcement learning, diffusion-based visuomotor policies, and foundation-model-based robotics that seek broader generalization and higher-level reasoning [1], [2], [6].

Imitation learning remains attractive because it avoids explicit reward design and allows a robot to learn directly from successful expert behavior. However, its limitations are well known: policies trained by behavior cloning often suffer from distribution shift, error accumulation, and poor recovery from unseen states. These issues become especially severe in contact-rich manipulation, where small prediction errors can lead to invalid grasps, unstable contact, or task failure. This motivates the transition toward methods that can move beyond direct imitation while preserving the useful structure contained in demonstrations [1], [5].

At the same time, the field is learning that capability alone is not enough. Seminar material from the course explicitly frames current robot learning as a move from "mimicking" toward reasoning, while also stressing that future success depends on combining learned policies with strong geometric, physical, and semantic priors. The same material highlights open challenges in generalization, robustness, and safety, which directly motivate the validation problem addressed in this report.

This report therefore argues that the main challenge in state-of-the-art robot learning is no longer only how to train more powerful policies, but how to validate those policies for safe, robust, and trustworthy deployment [3], [4], [7]. The remainder of the report first reviews robot learning from demonstration and modern imitation learning, then examines reinforcement learning and foundation-model robotics, and finally discusses validation, safety, and trustworthiness as central research bottlenecks.

II. From Learning from Demonstration to Modern Imitation Learning

Learning from Demonstration (LfD) has long been one of the most practical approaches to robot learning because it allows robots to acquire skills from expert behavior rather than handcrafted control rules. In robotic manipulation, demonstrations can encode rich information including motion trajectories, grasp choices, object relations, timing, and coordination patterns. This makes LfD especially appealing for tasks that are hard to model analytically. Seminar material in the course similarly emphasizes that demonstrations contain far more than low-level action labels and must be interpreted in terms of task-relevant structure.

A common realization of LfD is behavior cloning, in which a policy is trained in supervised fashion to map observations to expert actions. Its main strength is simplicity, but its main weakness is covariate shift: because the robot is trained mostly on successful expert states, even a small execution error may move it into unfamiliar states where mistakes compound over time. In robotics, this is not merely a statistical inconvenience but a control problem, since sequential errors can quickly cascade into collisions, loss of contact, or failure to complete the task [5].

A further limitation of classical imitation learning is that it often assumes the demonstrations already capture the most relevant parts of the task distribution. In practice, however, demonstrations are expensive to collect and frequently biased toward successful executions under relatively controlled conditions. This means that rare but important situations, such as partial occlusions, object slippage, or unexpected contact events, may be underrepresented in the training data. As a result, the learned policy may appear accurate during offline evaluation while still failing in deployment when confronted with disturbances that were not sufficiently covered during demonstration collection. This issue is especially important in robotics because real-world operation depends not only on average performance, but also on the ability to remain stable under uncertainty and recover from small deviations before they escalate into task failure [5].

Recent imitation learning research addresses these weaknesses by adopting richer policy representations. Diffusion-based policies, for example, model actions as samples from a learned distribution and iteratively refine action outputs during inference. This is scientifically important because robotic behavior is often multimodal: there may be several valid grasps or motion strategies for the same task, and a simple regression model can collapse these alternatives into an invalid "average" action. Diffusion Policy reports strong gains across multiple manipulation benchmarks and argues that the diffusion formulation handles multimodal action distributions and high-dimensional control particularly well [5].

Even with these advances, imitation learning remains bounded by the data and competence of the demonstrator. It is highly efficient as a way to transfer structure and initialize policies, but it does not by itself solve the broader problem of autonomous adaptation. For that reason, modern robot learning does not abandon imitation; rather, it treats imitation as a foundation that can be extended by optimization, interaction, and stronger model priors [1], [5].

III. Beyond Imitation: Reinforcement Learning and Policy Optimization

Reinforcement learning (RL) addresses one of the central limitations of imitation learning: a robot trained only to imitate cannot easily exceed the skill represented in its demonstrations [1]. In RL, the robot instead learns by interacting with the environment and optimizing a cumulative reward signal. This makes RL attractive in settings where demonstrations are incomplete, unavailable, or insufficient for covering the true range of task variation. In principle, RL enables autonomous improvement, recovery from off-distribution states, and discovery of behaviors not shown by an expert [1], [2].

In practice, however, robotics RL is difficult. Physical interaction is expensive, slow, and often safety-critical, while deep RL methods are known to be sample-inefficient and unstable without careful design. As a result, successful robotic RL usually relies on simulation, replay, action constraints, or hybrid formulations that combine demonstration priors with reward-based optimization. This is consistent with the seminar material, which explicitly contrasts the "foundation model" scaling path with an "efficient learning" path based on sample-efficient RL and IL algorithms, structured priors, and simulation [1], [2].

A further reason why RL alone remains difficult in robotics is that learning must take place under severe constraints on data efficiency, safety, and real-world interaction time. Unlike purely simulated benchmark settings, physical robots cannot explore freely at large scale without risking hardware wear, unsafe behavior, or long collection cycles. This makes naive trial-and-error learning impractical for many manipulation and service-robotics tasks. As a result, recent robot-learning systems rarely rely on unconstrained RL alone, and instead combine optimization with demonstrations, structured priors, or large pre-collected datasets in order to reduce the cost of exploration and improve stability during learning [1], [2].

Another important challenge in robotic reinforcement learning is reward design. In many manipulation and navigation tasks, the desired outcome can be described only at a high level, while the intermediate behaviors required for success are difficult to encode as a precise scalar objective. Poorly designed rewards may encourage unintended strategies, reward hacking, or brittle behavior that performs well in simulation but does not correspond to meaningful task completion in the real world. For this reason, many recent approaches reduce the burden on reward engineering by combining RL with demonstrations, curriculum learning, or prior knowledge about feasible motion and task structure [1], [2]. These strategies suggest that practical robot learning requires not only optimization power, but also careful alignment between the learning objective and the physical realities of robotic execution.

This observation leads to an important scientific conclusion: RL should not be understood as a full replacement for imitation learning, but as a complementary mechanism for extending imitation-based systems beyond the boundaries of expert data. Demonstrations provide strong priors and initialization, while RL supplies autonomous refinement, recovery, and optimization under novel conditions. The future of robot learning therefore appears increasingly hybrid, combining demonstrations, priors, simulation, and policy optimization rather than relying on a single paradigm [1].

IV. Foundation Models, Vision-Language-Action Models, and Generalist Policies

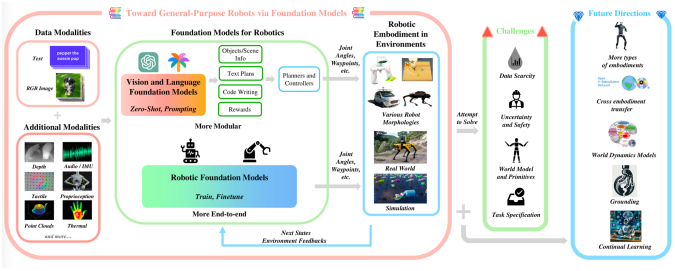

Fig. 1. Toward General-Purpose Robots via Foundation Models (Lecture 7 - Georgia Chalvatzaki)

A major recent shift in robot learning is the move from task-specific policies toward generalist models that aim to transfer across tasks, objects, instructions, and even embodiments. RT-1 introduced a transformer-based policy trained on large-scale real-world robotic data, arguing that generalization in robotics benefits from broad task-agnostic training and high-capacity architectures [1]. RT-2 extended this line of work by incorporating internet-scale vision-language pretraining directly into robotic control, representing actions as tokens and showing improved generalization to novel objects and semantically richer instructions [2]. Open X-Embodiment further pushed the scaling paradigm by combining data from 22 robot embodiments and 21 institutions to train cross-robot policies [6].

The emergence of such models also changes how researchers think about data in robotics. Earlier robot-learning systems were often trained on narrowly curated datasets collected for a single platform and task family, whereas generalist policies increasingly rely on heterogeneous data gathered across laboratories, embodiments, sensors, and control interfaces. This shift creates new opportunities for transfer, but it also introduces challenges in data consistency, action representation, and embodiment mismatch. A policy trained across different robots may learn more general semantic patterns, yet it must still translate those patterns into motions that are physically meaningful for a specific platform. Consequently, scalability in robot learning is not only a question of collecting more data, but also of designing representations that preserve cross-platform knowledge without losing embodiment-specific precision [2], [6].

These models are scientifically significant because they move robot learning from narrow trajectory imitation toward semantic grounding. Large language and vision-language backbones can help decompose user instructions, interpret open-vocabulary scenes, and connect abstract goals to low-level robotic behavior. Seminar material from the course describes this as the new paradigm of foundation models for robotics, where a single pretrained model is adapted across many robotic tasks instead of retraining from scratch for each task.

However, the same material also emphasizes the main limitation: such models still require grounding into physically actionable affordances. High-level semantic competence does not automatically imply physical precision, collision-free execution, or reliable contact handling. The seminar slides explicitly list generalization and robustness, sim-to-real-to-sim adaptation, and safety and predictability as the main grand challenges for foundation-model robotics.

For this reason, current evidence suggests that foundation models are best understood as an expansion of the robotics design space rather than a complete replacement for structured robot learning. Their strongest promise may lie in hybrid systems that combine pretrained semantic models with geometry-aware, physics-aware, and embodiment-aware priors. This hybrid view is also the one most strongly supported by the course material.

V. Validation, Safety, and Trustworthiness in Learned Robotic Systems

As robot learning becomes more capable, validation becomes more difficult. Traditional robotics often relies on manually specified logic, modular software components, and relatively explicit requirements. Learned policies, by contrast, are data-dependent, adaptive, and often opaque. High benchmark performance therefore does not automatically imply safe, predictable, or robust real-world behavior. This point is well supported in the software-engineering literature for autonomous robots, which argues that trustworthy deployment requires stronger development and assurance methods than those commonly used in laboratory prototypes [3], [4].

The course seminar material makes the same transition explicit: moving from single demonstrations, single tasks, and single robots in laboratory settings toward long-term deployment, diverse tasks, and interaction with groups of humans and robots requires new forms of evidence. It also states that evaluation is still costly, manual, and difficult to scale, partly because intended behavior is rarely captured in explicit, executable testing artefacts.

One promising answer is scenario-based validation. Ortega and colleagues propose composable and executable scenarios for simulation-based testing of mobile robots, motivated by the observation that developers struggle to manage the enormous variability of robotic testing scenarios. Their approach supports reusable scenario specification, transformation into simulation assets, and structured execution across variations [7]. The same ideas are reflected in the seminar slides, which show FloorPlan DSL, automated scenario generation, and the ExSce framework as mechanisms for scalable evaluation.

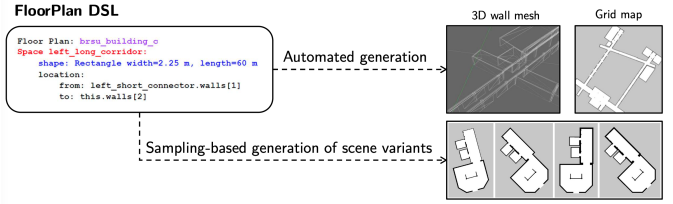

Fig. 2. FloorPlan DSL - Seminar slides (Lecture 6 - Nico Hochgeschwender)

Scenario-based testing also improves reproducibility because it makes evaluation conditions more explicit and reusable. Instead of relying on informal descriptions or one-off demonstrations, developers can represent relevant environmental structure, task assumptions, and expected operating conditions in a form that can be executed repeatedly across different experiments and system versions. Ortega et al. motivate this directly by identifying the variability of testing scenarios as a major obstacle for simulation-based testing in robotics and by proposing composable, executable scenarios as a way to manage that variability systematically. In this sense, scenario-based validation does not only help uncover corner cases, but also strengthens the reproducibility and comparability of empirical results in robot-learning research [4], [7].

An additional benefit of scenario-based validation is that it supports a clearer separation between task requirements and implementation details. Instead of evaluating a robotic system only through ad hoc end-to-end demonstrations, developers can specify relevant environmental properties, operating assumptions, and expected outcomes in a structured form. This makes it easier to compare systems, reproduce experiments, and identify which classes of situations lead to failure. For learned robotic policies, such structured evaluation is particularly valuable because it helps reveal whether generalization claims hold across meaningful variations in layout, timing, perception quality, and interaction context. In this sense, scenario-based testing does not only improve robustness analysis, but also strengthens the scientific quality of empirical evidence in robot learning [4], [7].

This matters because failures in learned robotic systems often arise in corner cases rather than average conditions. Systematic variation of scenarios enables developers to explore robustness across environment structure, task settings, and operational contexts. The seminar slides further connect this to generalized evaluation plans and stakeholder-aligned benchmarking, while also highlighting open challenges such as generating reliable oracles and deriving complete coverage from real-world data.

Safety and trustworthiness are therefore not properties of the learning algorithm alone. They emerge from the interaction between learned models, robot embodiment, software architecture, testing infrastructure, and operational assumptions. In that sense, validation is not a final step after training, but a central research problem that must evolve together with robot learning itself [3].

VI. Discussion & Conclusion

The development of robot learning beyond classical imitation learning reflects a broader shift in robotics from narrow task replication toward more general, adaptive, and semantically informed behavior. Imitation learning remains an important foundation because it enables efficient transfer of expert knowledge and preserves task structure without requiring complex reward design. However, its dependence on demonstration quality and its limited ability to recover from unseen states make it insufficient as a standalone solution for open-world robotic problems. Reinforcement learning extends this paradigm by enabling autonomous improvement through interaction, while foundation models and vision-language-action systems further broaden the scope of robot learning by linking perception, reasoning, and control. Together, these approaches significantly expand what robots can potentially do, but they also increase the complexity of failure modes, uncertainty, and deployment risks [1], [2], [5].

A central finding of this report is that progress in robot learning is no longer limited only by learning capability, but increasingly by validation and trustworthiness. More capable policies can generalize across a wider range of tasks and conditions, yet this same flexibility makes their behavior harder to predict, test, and guarantee. High benchmark performance does not necessarily imply safe or reliable performance in realistic environments, especially when learned policies interact with humans, uncertain scenes, or long-horizon tasks. For this reason, scenario-based testing, executable specifications, and systematic robustness evaluation are becoming essential components of modern robotic system design [4], [7]. These methods help ensure that claims of generalization are matched by meaningful evidence of dependable behavior under varying operating conditions.

At the same time, the literature reviewed in this report suggests that the future of robot learning will likely not be defined by a single dominant paradigm. Instead, the most promising direction is the development of hybrid systems that combine demonstrations, reinforcement learning, semantic foundation models, and structured priors derived from physics, geometry, and embodiment constraints. Such combinations make it possible to benefit from the efficiency of imitation, the adaptability of optimization-based learning, and the broad semantic knowledge provided by large-scale pretrained models. This hybrid perspective is particularly important in robotics, where successful behavior must remain grounded in the physical world rather than relying only on abstract predictive performance.

A further implication is that progress in robotics should be evaluated not only in terms of benchmark accuracy or task success rate, but also in terms of reproducibility, interpretability, and deployment readiness. As robot-learning systems become more complex, purely model-centric evaluation becomes insufficient. Scientific progress in this area will increasingly depend on whether researchers can provide convincing evidence that learned policies remain reliable across changing conditions, embodiments, and operational constraints. This shifts part of the research focus from isolated model performance toward the broader question of how capable robotic systems can be engineered as dependable systems.

In summary, moving beyond imitation is not only a question of building more powerful learning models, but also of establishing scientifically rigorous methods for evaluating and trusting their behavior in real-world deployment. The long-term success of state-of-the-art robotics will therefore depend on whether progress in capability can be matched by progress in validation, safety assurance, and system-level trustworthi-

ness. Only through this combination can robot learning move from impressive laboratory demonstrations to dependable use in practical environments [3].

VII. Acknowledgement

AI-based tools, including GPT models and integrated AI features in Overleaf, were used during the preparation of this report to support literature search and filtering. In addition, these tools were used to improve sentence formulation and overall wording in order to communicate the main ideas of the report more clearly and effectively. All sources were reviewed by the author, and the author is responsible for the final selection of references, the technical interpretation of the cited work, and the final written content of the report.

References

[1] A. Brohan, N. Brown, J. Carbajal, Y. Chebotar, J. Dabis, C. Finn, K. Gopalakrishnan, K. Hausman, A. Herzog, J. Hsu, J. Ibarz, B. Ichter, A. Irpan, T. Jackson, S. Jesmonth, N. J. Joshi, R. Julian, D. Kalashnikov, Y. Kuang, I. Leal, K.-H. Lee, S. Levine, Y. Lu, U. Malla, D. Manjunath, I. Mordatch, O. Nachum, C. Parada, J. Peralta, E. Perez, K. Pertsch, J. Quiambao, K. Rao, M. Ryoo, G. Salazar, P. Sanketi, K. Sayed, J. Singh, S. Sontakke, A. Stone, C. Tan, H. Tran, V. Vanhoucke, S. Vega, Q. Vuong, F. Xia, T. Xiao, P. Xu, S. Xu, T. Yu, and B. Zitkovich, "Rt-1: Robotics transformer for real-world control at scale," 2023.

[2] A. Brohan, N. Brown, J. Carbajal, Y. Chebotar, X. Chen, K. Choromanski, T. Ding, D. Driess, A. Dubey, C. Finn, P. Florence, C. Fu, M. Gonzalez Arenas, K. Gopalakrishnan, K. Han, K. Hausman, A. Herzog, J. Hsu, B. Ichter, A. Irpan, N. Joshi, R. Julian, D. Kalashnikov, Y. Kuang, I. Leal, L. Lee, T.-W. E. Lee, S. Levine, Y. Lu, H. Michalewski, I. Mordatch, K. Pertsch, K. Rao, K. Reymann, M. Ryoo, G. Salazar, P. Sanketi, P. Sermanet, J. Singh, A. Singh, R. Soricut, H. Tran, V. Vanhoucke, Q. Vuong, A. Wahid, S. Welker, P. Wohlhart, J. Wu, F. Xia, T. Xiao, P. Xu, S. Xu, T. Yu, and B. Zitkovich, "Rt-2: Vision-language-action models transfer web knowledge to robotic control," 2023.

[3] D. Brugali, A. Cavalcanti, N. Hochgeschwender, P. Pelliccione, and L. Rebelo, "Future directions in software engineering for autonomous robots: An agenda for trustworthiness [opinion]," *IEEE Robotics and Automation Magazine*, vol. 31, no. 3, pp. 186–204, 2024.

[4] H. Araujo, M. R. Mousavi, and M. Varshosaz, "Testing, validation, and verification of robotic and autonomous systems: A systematic review," *ACM Transactions on Software Engineering and Methodology*, vol. 32, no. 2, Mar. 2023.

[5] C. Chi, Z. Xu, S. Feng, E. Cousineau, Y. Du, B. Burchfiel, R. Tedrake, and S. Song, "Diffusion policy: Visuomotor policy learning via action diffusion," 2023.

[6] OXE Collaboration, "Open x-embodiment: Robotic learning datasets and rt-x models," 2023.

[7] A. Ortega, S. Parra, S. Schneider, and N. Hochgeschwender, "Composable and executable scenarios for simulation-based testing of mobile robots," *Frontiers in Robotics and AI*, vol. 11, p. 1363281, 2024.